Attention: *Upload via Canvas till 23:59 December 6 You can submit the solutions for problems 1,2 and 7 in a Jupyter Notebook file (with presented analytical solutions, code (Python 3 only) and explanations attached) and solutions for problems 3, 4, 5, 6 via separate pdf file. The file should be LaTeX-based, so no photo or scanned paper-based solutions are accepted.*

Attention 2: *I kindly remind you about collaboration policy. Indeed, it seems to me that receiving failure grade (F) after passing 5 weeks of the course is quite frustrating taking also into account that problems are relatively simple*

Problem 1

Points: 1

Dataset:

Test everything on artificial datasets, i.e. you can use:

```python
# data for censored fitting problem
import numpy as np
from numpy.random import randn

n = 20;  # dimension of x's
M = 25;  # number of non-censored data points
K = 100; # total number of points

np.random.seed(1)
c_true = randn(n)
X = randn(n, K)
y = np.dot(X.T, c_true) + 0.1 * np.sqrt(n) * randn(K)
print(c_true.shape)

# Reorder measurements, then censor
sort_ind = np.argsort(y)
X = X[:, sort_ind];
y = y[sort_ind[:M + 1]]
D = (y[M - 1]+y[M]) / 2
y = y[:M]
```

Task:

In some experiments there are two kinds of measurements or data available: the usual ones, in which you get a number (say), and *censored data*, in which you don't get the specific number, but you are told something about it, such as a lower bound. A classic example is a study of lifetimes of a set of subjects, say, laboratory mice. For those who have died by the end of data collection, we get the lifetime. For those who have not died by the end of data collection, we do not have the lifetime, but we do have a lower bound, i.e. the length of the study. These are the censored data values.

We wish to fit a set of data points,

$$\left((x^{(1)}, y^{(1)}), \ldots, (x^{(K)}, y^{(K)})\right)$$

where $x^{(i)} \in \mathbb{R}^n, y^{(i)} \in \mathbb{R}$ with a underlying linear model of the form $y \approx c^T x$. The vector $c \in \mathbb{R}^n$ is the model parameter, which we want to choose. We will use least-squares criterion, i.e. choose c via:

$$J = \sum_{k=1}^{K} \left(y^{(k)} - c^T x^{(k)}\right)^2 \longrightarrow \min_{c \in \mathbb{R}^n}$$

Here is the tricky part: some of the values of $y^{(k)}$ are censored: for those entries we only have a given lower bound. Assume $y^{(1)}, \dots, y^{(M)}$ to be given and $y^{(M+1)}, \dots, y^{(K)}$ to be censored, i.e. given a lower bound D. All the values of $x^{(k)}$ are assumed to be known.

- Explain how to find c and $y^{(M+1)}, \dots, y^{(K)}$ that minimize J.
- Implement the strategy on the data values from the given dataset. Report $\hat{c}$ (the model parameter found via this method). Also find $\hat{c}_{ls}$, the least-squares estimate of c obtained by simply ignoring the censored data samples, i.e. the least-squares estimate based on the data:

$$\Big((x^{(1)}, y^{(1)}), \dots, (x^{(M)}, y^{(M)})\Big)$$

The data file contains c_{true}, the value of c in the vector **c_true**. Use this to give the two relative errors:

$$\frac{\|c_{true} - \hat{c}\|_2}{\|c_{true}\|_2}, \; \frac{\|c_{true} - \hat{c}_{ls}\|_2}{\|c_{true}\|_2}$$

Problem 2

Points: 2

Dataset: veh_speed_sched_data.m

Task:

This is the same problem as problem 2 in the first assignmen but now you need to solve it in a different way. A vehicle (say, an airplane) travels along a fixed path of n segments, between $n+1$ waypoints labeled $0, \dots, n$. Segment i starts at waypoint $i-1$ and terminates at waypoint i. The vehicle starts at time $t = 0$ at waypoint 0. It travels over each segment at a constant (nonnegative) speed; s_i is the speed on segment i. We have lower and upper limits on the speeds: $s_{\min} \leq s \leq s_{\max}$. The vehicle does not stop at the waypoints; it simply proceeds to the next segment. The travel distance of segment i is di (which is positive), so the travel time over segment i is $\frac{di}{s_i}$. We let $\tau_i, i = 1, \dots, n$, denote the time at which the vehicle arrives at waypoint i. The vehicle is required to arrive at waypoint i, for $i = 1, \dots, n$, between times $\tau^i_{\min}$ and $\tau^i_{\max}$, which are given. The vehicle consumes fuel over segment i at a rate that depends on its speed $\Phi(s_i) = as_i^2 + bs_i + c$ kg/s.

You are given the data d (segment travel distances), $s_{\min}$ and $s_{\max}$ (speed bounds), $\tau_{\min}$ and $\tau_{\max}$ (waypoint arrival time bounds), and the the parameters a, b and c.

For the given form of the potentials, find the way to reduce the problem to a convex optimization problem and solve it using CVX (NB: you need not necessarily use one of the canonical convex optimization formulations we saw in the course). Use function step to plot speed vs time for the optimal schedule. What are relative pros and cons for using convex optimization vs. dynamic programming for such task?

Problem 3

Points: 1

Task:

Just to warm up consider the following SDP problem:

$$\begin{aligned} \underset{y \in \mathbb{R}^2}{\text{minimize}} \quad & 2y_1 \\ \text{subject to} \quad & \begin{pmatrix} 0 & 5y_1 & 0 \\ 5y_1 & 3y_2 & 0 \\ 0 & 0 & 4y_1 + 2 \end{pmatrix} \succeq 0 \end{aligned}$$

where $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ and inequality means positive-semidefiniteness of the corresponding matrix. It can be easily shown that even in this simple example strong duality does not hold. So, the task is the following: find the primal solution, derive a dual problem and obtain its solution, compare it with the primal solution and show that, in fact, strong dual fails to hold.

Problem 4

Points: 2

Task:

Consider the following function:

$$f_0(x) = \max_{i=1,\dots,m} \left(a_i^T x + b_i\right)$$

where $x \in \mathbb{R}^n$ and $a_i \in \mathbb{R}^n, b_i \in \mathbb{R}\ \forall i$. Consider the following convex optimization problem:

$$f_0(x) \longrightarrow \min_{x \in \mathbb{R}^n} \tag{1}$$

1. (0.5 points) Formulate an equivalent LP problem (by introducing new variable $t = \max\limits_{i=1,\dots,m} \left(a_i^T x + b_i\right)$) and form a dual of the corresponding LP problem.

2. (0.5 points) In fact, as one can notice, $f_0(x)$ is a convex function in x. Nevertheless, one may consider the following approximation for the objective function:

$$f_1(x) = \log\left(\sum_{i=1}^{m} e^{a_i^T x + b_i}\right)$$

   This function is convex as well as $f_0(x)$, but it is infinitely differentiable. Such approximation was also announced during lectures. As a generalization of such approximation method, one may consider the following function:

$$f_2(x) = \frac{1}{\alpha} \log\left(\sum_{i=1}^{m} e^{\alpha(a_i^T x + b_i)}\right)$$

   where $\alpha > 0$ is a scaling parameter, and the corresponding optimization problem:

$$f_2(x) \longrightarrow \min_{x \in \mathbb{R}^n} \tag{2}$$

   Reformulate the problem (2) via introducing new variables y_i and equality constraints $y_i = a_i^T x + b_i$ and form a dual problem of the new problem.

3. (1 point) Let x_0^* be the solution for the problem (1) and x_2^* be the solution for the problem (2). Show that:

$$0 \le f_2(x_2^*) - f_0(x_0^*) \le \frac{1}{\alpha} \log m \tag{3}$$

   This way, one deduces that $f_2(x_2^*)$ approaches $f_0(x_0^*)$ as α tends to infinity.

Hint: One side of (3) is trivial to show. To show the other side, consider the dual problems obtained in part 1 and part 2 of the exercise. Since both problems are convex, assuming that both programs are strictly feasible, strong duality holds. Since the duality gap is zero, consider the optimal solutions of the dual problems. Derive a relation between $f_2(x_2^*)$ and $f_0(x_0^*)$ and apply Jensen's inequality where necessary.

Problem 5

Points: 1

Task:

Consider equality-constrained LS problem:

$$\begin{aligned} \underset{x\in\mathbb{R}^n}{\text{minimize}} \quad & \|Ax-b\|_2^2 \\ \text{subject to} \quad & Cx = d \end{aligned}$$

where $A \in \mathbb{R}^{m\times n} : \text{rk}(A) = n, C \in \mathbb{R}^{p\times n} : \text{rk}(C) = p$. Write down the KKT conditions and derive expressions for optimal primal and dual solutions.

Problem 6

Points: 2

Task:

Consider the following LP:

$$\begin{aligned} \underset{x\in\mathbb{R}^n}{\text{minimize}} \quad & c^T x \\ \text{subject to} \quad & Ax = b \\ & x \geq 0 \end{aligned} \tag{4}$$

where the inequality is, obviously, element-wise and similar optimization problem:

$$\begin{aligned} \underset{x\in\mathbb{R}^n}{\text{minimize}} \quad & c^T x - \tau \sum_{i=1}^{n} \ln(x_i) \\ \text{subject to} \quad & Ax = b \end{aligned} \tag{5}$$

where $\tau > 0$ is a parameter. It is clear that as any of x_i's tends to zero, the corresponding log-term will tend to negative infinity and this way make the objective function tend to positive infinity. Assuming that the primal LP and its dual are both strictly feasible: sets $\{x : x > 0, Ax = b\}$ and $\{y : A^T y + c > 0\}$ are non-empty, answer the following questions:

1. (0.5 points) Derive the dual of (4) and write down KKT conditions.
2. (0.75 point) Derive the dual of (5) and write down KKT conditions.
3. (0.75 points) Compare two cases. What happens when parameter τ is taken large?

Problem 7

Points: 2

Task:

Consider the following convex quadratic program:

$$\begin{aligned} \underset{x_1,x_2\in\mathbb{R}}{\text{minimize}} \quad & 2x_1^2 + 2x_2^2 - x_1 x_2 \\ \text{subject to} \quad & 2x_1 + 3x_2 \geq 1 \\ & 5x_1 + 4x_2 \geq 1 \end{aligned}$$

Derive the dual program. Put both programs into CVX and solve them. Verify that the optimal values are the same. State the KKT conditions. Check that the optimal primal and dual variables satisfy them.